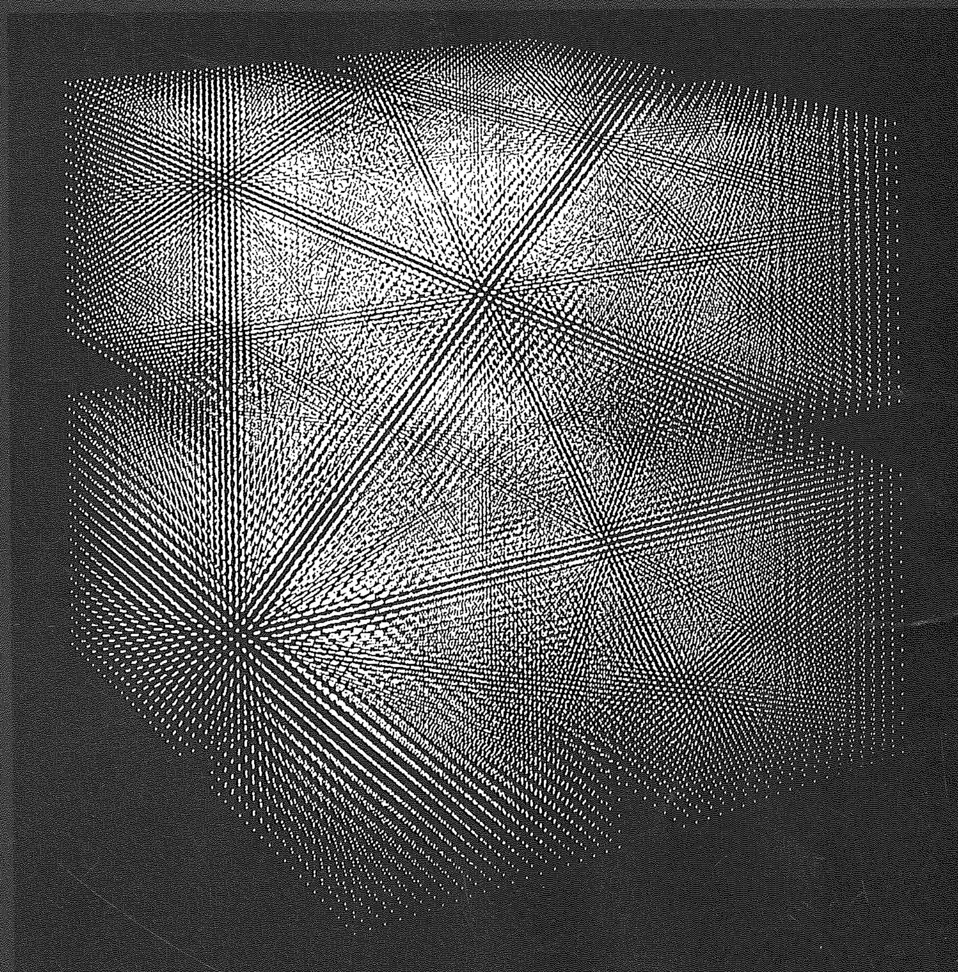
The Connection Machine Family.



A new standard in large scale computing.

Connection Machine® Family: Parallel Computers that are Easy to Program.

Thinking Machines Corporation produces a family of high performance computer systems. The largest member of the family is the 64,000 processor CM-2 with performance in excess of 2500 Mips, and floating point performance above 2500 MFlops. The systems are programmed through familiar programming environments, like UNIX, using parallel extensions of conventional languages, like Fortran, Lisp, and C. Connection Machine systems are currently being used in a range of applications including data base retrieval, image processing, computer aided design and floating-point intensive scientific calculations.

The Connection Machine system is easier to program than other parallel computers because it supports a data parallel style of computation. This style allows users to express a parallel computation in a natural way, operating on all the data at once, without worrying about issues like partitioning and synchronization. The machine's general communications network automatically adapts to the structure of the application.

The Connection Machine computer contains 64,000 processors connected by a general communication network. The air-cooled system presents no special environmental requirements.

The DataVault™ mass storage system provides 10 Gbytes of fault tolerant mass storage connected to one of the Connection Machine high bandwidth I/O channels.

The high resolution graphics display loads from Connection Machine memory at 1 Gbit per second.

Data Parallel Computing:

Looking at the whole problem at once means computing on every element of a data structure at the same time. The Connection Machine system's tens of thousands of processors allow it to do this in a straightforward way: by attaching a separate processor to each element of a data structure.

The type of data structure depends on the application. In many language processing applications, data parallelism means a processor for every word or every meaning. For data base applications, it means a processor for every document. In a numeric simulation, it may mean a processor for every element of a matrix. And in image processing, data parallelism means a processor for every picture element, or pixel, in an image.

The performance advantages of data parallelism are dramatic. The parallelism inherent in the data structures grows with the size of the data. Processing all the pixels in an image, for example, is as fast as processing one pixel, because they can all be computed at the same time. A whole data base can be searched in the time it takes to search one document. When computing the flow of air over an airplane wing, the Connection Machine system calculates the flow over all parts of the wing simultaneously, just as it works in reality. Speedups of 1000 or more are common with data parallelism.

If data parallelism is so simple and powerful, why don't all big computers work this way? Because in order to make these programs work, tens of thousands of processors have to work together. They have to communicate. The Connection Machine system solves the communications problem in hardware so the natural algorithms work well.

What makes the Connection Machine system work are the connections. Inside the machine there is a very flexible high-bandwidth communication network that moves data between processors at billions of bits per second. Routing circuits on every chip automatically steer data along the fastest paths, helping to make programming simpler. There is no need to adapt your application to the structure of a fixed architecture like a grid, ring, hypercube, or tree. Instead, the Connection Machine system adapts to your application, by dynamically forming the connections that are needed.

CM-2 Specifications

Typical Application Performance* (fixed point)

General Computing	2500 Mips
Terrain Mapping	1000 Mips
Document Search	6000 Mips

Interprocessor Communication (32-bit messages)

Regular Pattern	250 million per second
Random Pattern	80 million per second
Sort 65,536 32-bit keys	30 milliseconds

Variable Precision Fixed Point

64-bit integer add	1500 Mips
32-bit integer add	2500 Mips
16-bit integer add	3300 Mips
8-bit integer add	4000 Mips
64-bit move	2000 Mips
32-bit-move	3000 Mips
16-bit move	3800 Mips
8-bit move	4500 Mips

Double Precision Floating Point

Average (4K x 4K matrix multiply)	2500 MFlops
Dot product	5000 MFlops

Single Precision Floating Point

Average (4K x 4K matrix multiply)	3500 MFlops
Dot product	10,000 MFlops

General Specifications

Processors	65,536
Memory	512 Mbytes
Memory Bandwidth	300 Gbits per second

Input/Output Channels

Number of Channels	8
Capacity per Channel	40 Mbytes per second
Maximum Transfer Rate	320 Mbytes per second

Physical Dimensions

Size	56″ x 56″ x 62″
Weight	2,600 lbs.

Environmental Requirements (does not include host)

Power Dissipation	28 KW
Power Input	Four 30-amp 3-phase 110/208v
Operating Temperature	70°F ± 5°F
Operating Relative Humidity	50% ± 10%

*Thinking Machines Corporation believes all specifications are accurate as of the date of publication. Thinking Machines Corporation cannot, however, be responsible for inadvertent errors. Product specifications are subject to change without notice. For further detail, see the Product Specification Sheet. Mips = Millions of instructions per second; MFlops = Millions of floating point operations per second.

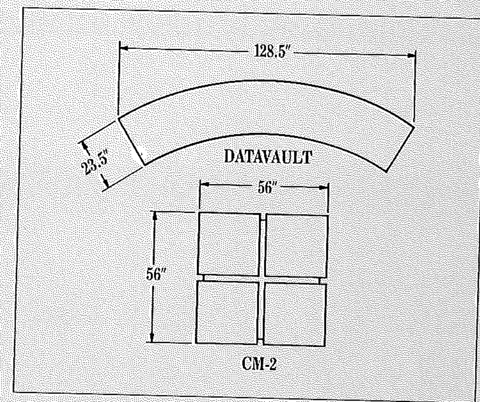
System Components

Programming Environment:
The user interacts with the Connection Machine system through a conventional front end system, such as a Digital VAX[1] or a Symbolics 3600.[2] The system is programmed via the front end, using familiar editors and utilities. File structures and network protocols are supported there as well, as are the full range of standard VAX and Symbolics peripherals.

Languages:
C* supports the data parallel programming style while making minimal extensions to the C language itself. One new parallel data type is introduced.

Lisp is supported at two levels. CM-Lisp offers fully automatic allocation of parallel data structures. *Lisp allows users to get closer to the hardware. Both CM-Lisp and *Lisp are extensions of Common Lisp.

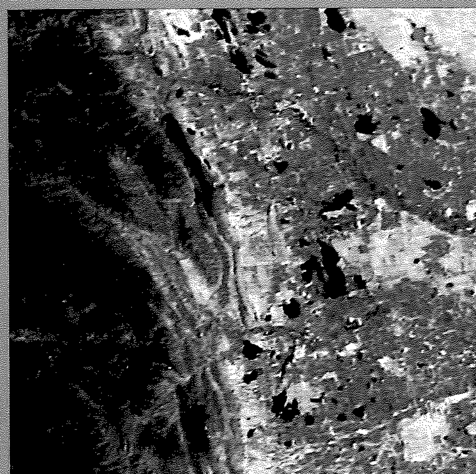
Fortran supports the data parallel programming style by using Fortran 77 with vector and control extensions meeting the emerging standard of Fortran 8x.



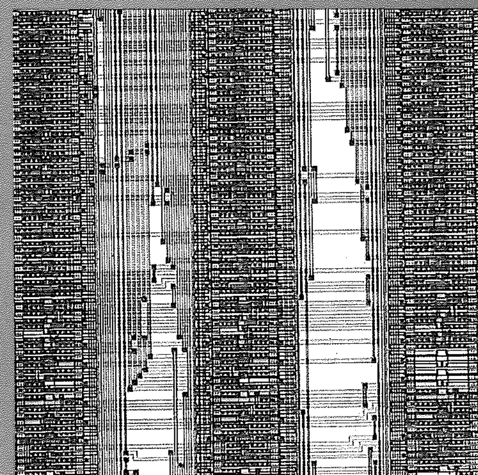
Mass Storage:
DataVault provides highly fault tolerant storage for Connection Machine data. DataVaults are available in 5 or 10 Gbytes with a 40 Mbyte per second transfer rate. Up to eight DataVaults may be used in parallel, for a total transfer rate of 320 Mbytes per second. Data is redundantly stored across multiple disk units so that any single unit can fail without loss of information.

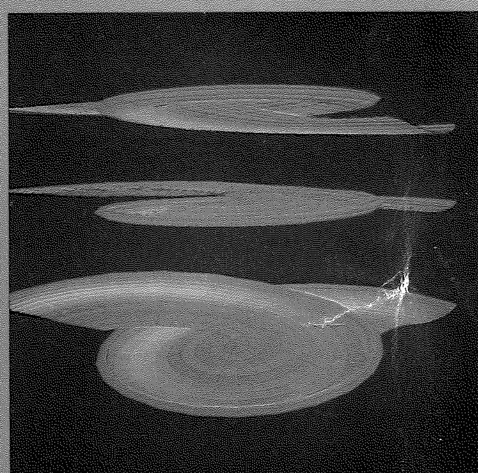
Graphics Display:
The system's high resolution graphic display can be loaded from Connection Machine memory at 1 Gbit per second, fast enough for real time animation. The display stores a 1280 x 1024 color image, with 24 bit planes.

Terrain Mapping

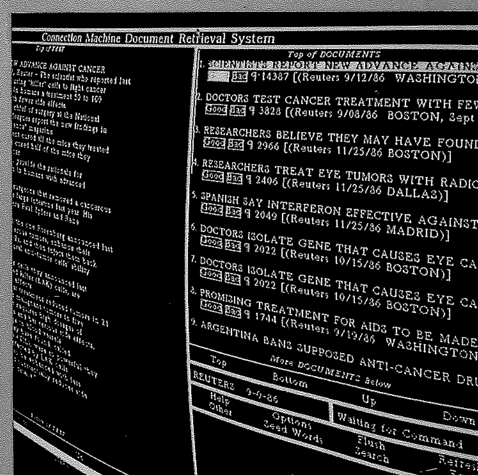


VLSI Simulation





Fluid Dynamics



Document Retrieval

Thinking Machines Corporation
245 First Street
Cambridge, MA 02142-1214
(617) 876-1111